JOURNAL OF THE EXPERIMENTAL ANALYSIS OF BEHAVIOR 2001, **76,** 95–114 NUMBER 1 (JULY)

PERCEPTUAL CLASSES ESTABLISHED WITH FORCED-CHOICE PRIMARY GENERALIZATION TESTS AND TRANSFER OF FUNCTION

KENNETH F. REEVE AND LANNY FIELDS

QUEENS COLLEGE AND
THE GRADUATE SCHOOL AND UNIVERSITY CENTER OF
THE CITY UNIVERSITY OF NEW YORK

In Experiment 1, 20 college students learned two identity conditional discriminations using squares that differed in interior-fill percentage (called Fill23 and Fill77). A two-choice generalization test was then presented with number of test trials varied across groups of subjects. The test samples were 19 squares that ranged in fill value from 23% to 77%; the comparisons were squares with Fill23 and Fill77. The resulting gradients did not vary as a function of number of test trials. When the generalization test was repeated with a third comparison, ''neither,'' the ranges of fill values that occasioned the exclusive selection of Fill23 or Fill77 were direct functions of the number of prior two-choice generalization trials. Finally, a discriminability test revealed that Fill23 and Fill77 were discriminable from the intermediate fill values. In Experiment 2, perceptual classes were established with 5 new students using 760 forced-choice generalization test trials. The students were then trained to select a different glyph in the presence of Fill23 and Fill77, followed by a three-choice generalization test in which the 19 fill stimuli served as samples and the two glyphs served as comparisons. The gradients overlapped with those previously obtained during the three-choice generalization test that used Fill23 and Fill77 as comparisons. Finally, a discriminability test showed that many adjacent stimuli along the fill dimension were discriminable from each other. Together, the results of both experiments suggest that ranges of fill-based stimuli functioned as members of perceptual classes, and each class also functioned as a transfer network for a new selection-based response.

Key words: perceptual classes, generalization, discriminability, natural categories, keyboard press, adult humans

---

The stimuli comprising a perceptual class can be arrayed along a simple physical dimension, such as line length (Fields, Reeve, Adams, Brown, & Verhave, 1997) or sound frequency (Njegovan, Ito, Mewhort, & Weisman, 1995; Risley, 1964), or along a complex mathematically defined dimension, such as ''compactness'' of geometric shapes (Hrycenko & Harwood, 1980). Alternatively, the dimension along which such stimuli are arrayed can be defined psychometrically. Such a dimension could be created by sorting the stimuli in a potential class according to their ''typicality'' (Bourne, Dominowski, & Loftus, 1979; Cook, Wright, & Kendrick, 1990; Lea & Ryan, 1984; Rosch & Mervis, 1975). For example, when presented with different exemplars of tree images, human participants may be instructed to rank the typicality of each exemplar. Such perceptual classes have been referred to as basic level (Rosch & Mervis, 1975; Rosch, Mervis, Gray, Johnson, & Boyes-Bream, 1976), natural (Herrnstein, 1990; Lea & Ryan, 1984; Wasserman & DeVolder, 1993; Wittgenstein, 1968), ill-defined (Bourne et al., 1979; Homa & Little, 1985; Neisser, 1967; Omohundro, 1981), or probabilistic categories (Medin & Smith, 1984).

To infer that stimuli are functioning as members of a perceptual class, two functional properties should be demonstrated. First, any response emitted in the presence of a narrow range of stimuli must also be emitted, without direct training, in the presence of a much larger range of novel stimuli during a generalization test (Adams, Fields, & Verhave,

---

Financial support for this research was provided by U.S. Army Research Institute Contracts MDA903-90-C-0132 and DASW01-96-K-0009. Portions of these data appeared in a doctoral dissertation submitted by the first author to the Graduate School and University Center of the City University of New York. Portions of these data were also presented in the symposium ''Variables and Processes That Influence the Emergence of Perceptual Classes and Equivalence Classes'' at the annual convention of the Association for Behavior Analysis, Orlando, Florida, May 1998. We thank all the Queens College students who participated.

Correspondence concerning this article should be sent to Kenneth F. Reeve, Department of Psychology, Caldwell College, 9 Ryerson Ave., Caldwell, New Jersey 07006 (E-mail: DrKFReeve@aol.com) or Lanny Fields, Department of Psychology, Queens College/CUNY, 65-30 Kissena Boulevard, Flushing, New York 11367 (E-mail: Lanny_Fields@qc.edu).

1993; Bhatt, Wasserman, Reynolds, & Knauss, 1988; Cook et al., 1990; Fields & Reeve, 2000; Fields et al., 1997; Herrnstein, 1990; Keller & Schoenfeld, 1950; Lea, 1984; Wasserman, Kiedinger, & Bhatt, 1988; Zentall, Jackson-Smith, & Jagielo, 1990). Second, many of the stimuli to which the response generalizes must be discriminable from each other (Cook et al., 1990; Fields & Reeve, 2000; Fields et al., 1997; Keller & Schoenfeld, 1950; Lea, 1984; Wasserman et al., 1988). The latter property is necessary to conclude that generalization among the stimuli comprising the perceptual class is not due to an inability to discriminate among the stimuli. If this were the case, referring to this set of many indistinguishable stimuli as a ''class'' would be trivial because responding would essentially be under the control of a single stimulus value.

Once established, an additional property of the class may be demonstrated. Specifically, it should serve as a transfer network for new behavioral functions. This can occur in any of three different ways: (a) A new response trained to occur in the presence of one class member should also occur in the presence of the remaining class members with no additional training, (b) a new stimulus linked to a class member through conditional discrimination training should be selected in the presence of other class members, or (c) a new discriminative function acquired by one class member should also be occasioned by the other members (Dougher, Augustson, Markham, Greenway, & Wulfert, 1994; Dougher & Markham, 1994; Fields, Adams, Buffington, Yang, & Verhave, 1996; Hayes, 1991; Sidman, Wynne, Maguire, & Barnes, 1989).

To investigate the establishment of perceptual classes, most researchers have used simple discrimination training procedures in which many stimuli are drawn from two or more potential classes (e.g., Bhatt et al., 1988; Cook et al., 1990; Engelmann & Carnine, 1982; Herrnstein, 1990; Herrnstein, Loveland, & Cable, 1976; Honig & Stewart, 1988; Lubow, 1974; Malott & Siddall, 1972; Porter & Neuringer, 1985; Wasserman et al., 1988). In contrast, conditional discrimination training procedures have been used by a smaller number of researchers to investigate the establishment of perceptual classes. Typically, single exemplars drawn from different potential classes are used during conditional discrimination training. In both methods, subsequent generalization tests are conducted by presenting many novel stimuli that can be arrayed along the same dimension as the stimuli used in training (Fields, Adams, Brown, & Verhave, 1993; Fields et al., 1996, 1997; Fields, Reeve, Adams, & Verhave, 1991; Njegovan et al., 1995; Porter & Neuringer, 1985; Wasserman et al., 1988).

Fields et al. (1997), for example, exposed college students to identity conditional discrimination training using two lines that were 1 and 25 units in length. This was followed by a 250-trial forced-choice generalization test in which lines of intermediate length were presented as samples with the 1- and 25-unit lines as comparisons. The resulting generalization gradients showed that a range of short and long test lines always occasioned the selection of the 1-unit and 25-unit lines, respectively. Thus, these ranges may have been functioning as two perceptual classes (Fields & Reeve, 2000). As the sample test lines became more intermediate in length, the choice of a given comparison decreased systematically. These intermediate-length lines, then, were not functioning as members of either the long- or short-line perceptual class. Similar results have been obtained by other researchers who used forced-choice testing to investigate responding in the presence of stimuli arrayed along single stimulus dimensions (Cross & Lane, 1962; Migler & Millenson, 1969; Risley, 1964).

As noted by Fields et al. (1997), one shortcoming with forced-choice conditional discrimination trials is that they do not allow independence of classes arrayed along the same dimension. That is, selection of one comparison is always the complement of the other. During the forced-choice generalization tests in the Fields et al. study, students were required to assign each sample line to one of the two comparisons, even if the sample did not appear to be perceptually similar to either comparison. Thus, selection responses occasioned by given test lines may have been based on perceptual class membership or the forced-choice nature of the trial format. To separate these sources of control, Fields et al. repeated the prior two-choice generalization test with the addition of a third response option, called a ''neither'' or default comparison (Innis, Lane, Miller, & Critchfield, 1998; Roche & Barnes, 1996).

During this three-choice generalization test, some of the intermediate lines that previously occasioned selection of either the 1- or 25-unit lines now occasioned selection of the neither comparison. Because the likelihood of selecting the 1-unit line in the presence of the test lines was no longer the complement of selecting the 25-unit line, the inclusion of the neither comparison resulted in the separation of the dimension into two regions, each of which functioned as an independently defined perceptual class (Fields et al., 1993, 1997; Innis et al., 1998).

The results of the Fields et al. (1997) study raise the following question: What variables were responsible for the establishment of the two independent perceptual classes measured during three-choice generalization testing? Because identity training and forced-choice generalization testing always preceded three-choice testing, it is possible that both variables, alone or in combination, were responsible for the formation of perceptual classes. If identity training alone was sufficient to establish the classes, the subsequent forced-choice testing would have no effect on the likelihood of class establishment or on the range of stimuli that functioned as members of each class. Because all students were exposed to forced-choice generalization testing in the Fields et al. study, however, this could not be assessed.

Alternatively, forced-choice testing may have induced class formation. It may be that unreinforced selections, occasioned by the conditional discrimination test trials, established stimulus–stimulus relations between the endpoints of the dimension and ranges of short and long lines. Indeed, various studies have demonstrated the establishment of such stimulus–stimulus relations without the use of reinforcement. Specifically, these studies have demonstrated the emergence of unreinforced conditional selections in the presence of new discrimination trials (Saunders, Saunders, Kirby, & Spradlin, 1988), the use of multiple negative comparison training to establish new conditional discriminations (Adams, Fields, & Verhave, 1999), and the use of sensory preconditioning or respondent-type stimulus pairings to establish stimulus relations (Leader, Barnes, & Smeets, 1996; Leader, Barnes-Holmes, & Smeets, 2000).

In addition, if unreinforced conditional selections were responsible for the establishment of the stimulus–stimulus relations that led to class formation in the Fields et al. (1997) study, would the number of such trials affect the size of the resultant classes? A substantial literature suggests that the strength of such stimulus–stimulus relations is a function of the number of training trials (e.g., see Hearst, 1988). This would lead one to predict a relation between the range of stimuli functioning as members of a perceptual class and the number of prior forced-choice generalization test trials presented immediately following identity training. To date, however, no studies have isolated the effects of conditional discrimination identity training and forced-choice generalization testing on the formation of perceptual classes. The purpose of Experiment 1 was to make such a determination.

In Experiment 1, the potential perceptual classes consisted of squares containing different percentages of "fill." All students first received single-exemplar identity training. Some students then received different numbers of forced-choice generalization test trials. Other students received no forced-choice generalization testing. Following this, a three-choice generalization test that included a third response option, "neither," was presented. Performances during the three-choice test determined whether the number of prior forced-choice generalization test trials affected (a) the likelihood of perceptual class formation and (b) the size of the perceptual classes established. Performances during the three-choice generalization test were examined because the inclusion of the neither comparison allowed a measure of two independent perceptual classes in which membership of a stimulus in one class did not reduce the size of the other class (Fields et al., 1993, 1997; Innis et al., 1998).

EXPERIMENT 1

METHOD

Participants

Twenty undergraduate introductory psychology students at Queens College were randomly assigned to four different experimental groups. No student was familiar with the research area. Students received partial course credit upon completion of the experiment independent of performance. For each student, the experiment was completed in a single session that lasted approximately 3 to 4 hr.

Apparatus and Stimuli

The experiment was conducted with an IBM®-compatible personal computer that displayed all stimuli on a 36-cm black-and-white monitor. Responses consisted of touching specific keys on a standard keyboard. The experiment was controlled by custom software that programmed all stimulus presentations and recorded all keyboard responses.

Figure 1 depicts 5 of the 19 stimuli used in Experiment 1. Each stimulus was a borderless square (5 cm by 5 cm) that contained a different percentage of white pixels on a black background. The percentage of white pixels defined each square's fill value, which ranged from 23% to 77% (Fill23 to Fill77) in 3% increments. For each fill value, the pattern of pixels was randomly generated preexperimentally and remained fixed throughout the experiment. Fill23 and Fill77 defined the endpoints of the dimension of fill percentage. Fill50 was the physical midpoint of the dimension. Fill23 and Fill77 served as comparisons on all trials. During primary generalization tests, each of the 19 fill values, including the endpoint values, was presented as a sample.

Procedure

Trial format, contingencies, and responses within a trial. All trials used a matching-to-sample format. A trial began when "Press ENTER" appeared on the screen. Pressing the enter key cleared the screen and displayed a sample stimulus at the top center of the monitor. Pressing the space bar displayed two comparison stimuli (Fill23 and Fill77) at the bottom left and right corners of the screen along with the sample. During trials in which the third comparison was programmed, the words "If NEITHER press 4" appeared between the Fill23 and Fill77 comparisons.

During a trial, the left or right comparison was selected by pressing the 1 or 2 key, respectively. Pressing the 4 key selected the nei-

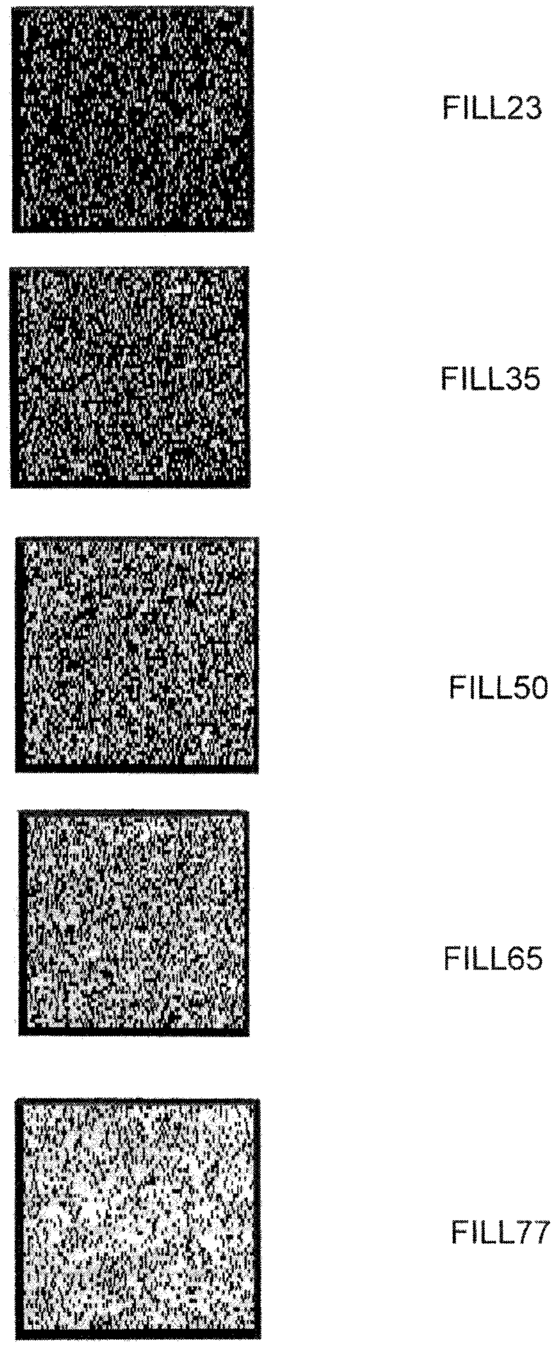


Fig. 1. Five of the 19 stimuli used in Experiment 1. The value following Fill represents the percentage of filled pixels for that stimulus. Fill23 and Fill77 were the endpoints of the dimension.

ther comparison, when available. A comparison selection cleared the screen and concurrently displayed a feedback message centered on the screen. When informative feedback was scheduled, the messages ''RIGHT'' or ''WRONG'' appeared, depending on the accuracy of the comparison selection. The message remained on the screen until the R (for ''RIGHT'') or W (for ''WRONG'') key was pressed. During some training and all test trials, noninformative feedback (a dashed line) was scheduled following a comparison selection to signal the end of a trial. The dashed line remained on screen until the student pressed the E key to end the trial. After the appropriate observing response was made (pressing the R, W, or E key), the screen was cleared and the next trial began.

Trial block structure and feedback contingencies. Each phase of training and testing was conducted with blocks of trials. Within all blocks, trials were presented in a random order without replacement. At the start of training, a block was presented repeatedly with informative feedback after each comparison selection until all trials within the block occasioned 100% correct responding. Thereafter, the percentage of trials that occasioned informative feedback was reduced to 75%, 25%, and finally to 0% over successive blocks as long as 100% accuracy within a block was maintained. During feedback reduction, the trials that were followed by informative feedback were randomly determined. If 100% correct responding was not achieved within three blocks at a given feedback level during training, the student was returned to the previous feedback level for that particular block.

Phase 1: Instructions and keyboard familiarization. Prior to the experiment, students were presented with the following instructions on the screen:

> Thank you for volunteering to participate. Please do not touch any keys on the keyboard yet! You will be presented with many trials. Each trial contains three or four CUES consisting of shapes, symbols, or common words. Your task is to discover how to respond correctly to the cues by pressing certain keys on the computer's keyboard. Initially, INSTRUCTIONS will tell you how to respond to the cues, and LABELS will help you identify the cues on the screen. The labels and instructions will slowly disappear. The experiment will be conducted in phases. When each phase ends, the computer will sometimes tell you how you did. When ready, please press the space bar to continue.

After pressing the space bar, students learned to emit the appropriate keyboard responses to complete a trial. To accomplish this, 16 trials, each containing three English words such as KING, QUEEN, and CAMEL, were presented. The semantic relatedness between the sample word (e.g., KING) and one of the comparisons (e.g., QUEEN) was used to prompt the selection of the correct comparison. Informative feedback (the words RIGHT or WRONG) followed each comparison selection (refer to Fields et al., 1997, for further details).

Correct responding to the stimuli in a trial during Phase 1 was also facilitated by instructional prompts (e.g., ''Make your choice by pressing key 1 or 2'') that were deleted in a serial manner across trials (refer to Fields et al., 1997, for further details). Phase 1 ended when performance exceeded 85% accuracy (14 of 16 correct trials) during a single block with no prompts. For the remainder of the experiment, if a nonexperimentally defined key was pressed during a trial, the instruction used to prompt the appropriate key press during keyboard familiarization (Phase 1) reappeared on the screen for three subsequent trials.

Phase 2: Two-choice identity conditional discrimination training. Students received identity conditional discrimination training in which Fill23 and Fill77 were randomly alternated as samples. The two comparisons, Fill23 and Fill77, each appeared equally often in random order on the left and right of the screen. A correct response was the selection of the comparison that was identical to the sample. When 100% informative feedback was scheduled, each block contained 32 trials. When 75%, 25%, or 0% feedback was scheduled, each block contained 16 trials. Feedback was reduced as previously described. Phase 2 was completed when no errors were emitted during a block with 0% informative feedback.

Phase 3: Two-choice primary generalization testing. Once the identity conditional discriminations were established, a primary generalization test for fill percentage was conducted for 15 of the 20 students. Within a test block, each of the 19 fill-value samples was presented twice in random order for a total of 38 trials. Fill23 and Fill77 were the two comparisons. For each sample, a given comparison appeared once on the left and once on the right in random order. All comparison selections were followed by noninformative feedback.

The number of two-choice generalization trials differed across the four groups. Students received either 152 (4 blocks), 456 (12 blocks), 760 (20 blocks), or 0 two-choice generalization trials. This assessed the potential effects of the number of two-choice generalization test trials on the likelihood of subsequent perceptual class formation and on the size of the resultant classes.

Phase 4: Training use of the neither comparison. In Phase 4, students learned how to select a default-response option (the neither comparison). To accomplish this, a 12-trial training block was used in which the sample and two of the three comparisons were English words (e.g., SOAP, COMPUTER, TRASH). The third comparison, the word NEITHER, was located between the two other comparisons on all trials. During trials in which one of the word comparisons was semantically related to the sample, its selection was followed by the feedback message RIGHT. During trials in which neither of the two comparison words was related to the sample, selection of the neither comparison was followed by RIGHT. When responding in the block reached 100% accuracy, feedback was reduced as previously described. Phase 4 ended when a student emitted no errors during a block with 0% informative feedback (refer to Fields et al., 1997, for further details). The neither comparison was then included during all subsequent phases in Experiment 1 to allow students to assign a given sample fill value to neither the Fill23 nor the Fill77 comparison.

Phase 5: Three-choice primary generalization testing. Primary generalization gradients were obtained by presenting the same generalization test block used in Phase 3 with the addition of the neither comparison on all trials. Inclusion of the neither comparison allowed an independent measure of the range of low- and high-fill stimuli that occasioned the exclusive selection of the Fill23 and Fill77 comparisons, respectively. The three-choice generalization test block was presented four times (152 trials) to all students.

Phase 6: Discriminability training and testing. To assess the discriminability of the endpoint stimuli (Fill23 and Fill77) from the fill values used during generalization testing (Fill26 through Fill74), students were presented with the same trial block used in the three-choice generalization test in Phase 5. At the start of Phase 6, however, all comparison selections were followed by informative feedback. When Fill23 or Fill77 was the sample, selection of the identical comparison was correct. When the remaining 17 intermediate fill values (26% to 74%) were presented as samples, selection of the neither comparison was correct. Once 100% correct responding was obtained, feedback was then reduced to 75%, 25%, and 0% of the trials, as previously described. Performances during the last four blocks were used to determine the degree of discriminability between the fill values used in training and testing.

RESULTS

Identity conditional discrimination performances. A minimum of 80 trials was required to complete identity conditional discrimination training. Students learned the identity conditional discriminations within a mean of 99.2 trials. There was no systematic difference in the number of trials to reach criterion across students in all groups. These results were confirmed statistically, $F(3, 16) = 0.42$, *ns.* Thus, differences in test performances during subsequent experimental conditions could not be attributed to differences in student performances during identity training.

Measuring the size of perceptual classes. The purpose of Experiment 1 was to determine how the number of forced-choice test trials affected both the likelihood of perceptual class formation and the size of the resultant classes. These effects were indexed by the range of low- or high-fill stimuli that occasioned the nearly exclusive selection of Fill23 or Fill77, respectively, during the three-choice generalization tests. To quantify the range indexing the size of the high-fill class, the value of the lowest fill stimulus (e.g., Fill65) that occasioned selection of Fill77 on at least 87.5% of the trials was subtracted from the endpoint value (Fill77). In this example, the range would be 12% ($77 - 65 = 12$). To

Table 1

Percentage of trials in which Fill23 was selected during two-choice primary generalization tests for individual students in Experiment 1.

Group	Student	Sample fill values (%)																		
		23	26	29	32	35	38	41	44	47	50	53	56	59	62	65	68	71	74	77
152	EK	100	100	100	88	75	100	88	75	38	25	—	—	—	—	—	13	—	—	—
	IM	100	100	100	100	100	88	88	100	50	25	13	13	—	—	—	—	—	13	—
	GL	100	100	100	88	100	100	100	100	100	100	75	50	25	38	13	—	—	—	13
	EC	100	100	88	100	100	88	75	100	75	63	25	13	—	—	—	13	—	—	—
	JU	100	100	100	100	100	63	88	75	75	75	38	—	—	—	—	13	—	13	—
	M	100	100	98	95	95	88	88	90	75	58	30	15	5	8	3	8	0	5	3
456	PF	100	100	96	96	96	96	96	75	67	67	21	13	4	—	4	—	—	—	—
	MD	96	96	96	100	100	96	75	92	42	13	—	—	—	—	4	—	—	—	4
	IS	92	100	100	96	100	100	100	88	75	63	29	4	8	—	—	—	4	4	—
	DV	100	100	100	100	100	100	100	88	75	68	—	—	—	—	4	—	4	—	—
	DL	100	100	100	100	100	88	88	71	54	25	8	—	4	—	—	—	—	—	—
	M	98	99	98	98	99	96	92	83	63	45	12	3	3	0	3	0	2	1	1
760	KS	100	100	100	100	95	93	78	53	23	35	8	3	3	—	—	-	—	—	—
	PP	100	100	95	98	100	93	95	90	75	85	40	13	8	3	3	13	13	3	5
	DJ	100	95	100	100	98	100	100	90	58	60	25	10	8	—	—	5	—	—	3
	AT	98	95	95	97	95	85	80	60	15	23	5	3	—	—	—	5	—	3	—
	NR	95	93	93	93	93	90	78	63	35	35	15	5	—	—	—	3	3	—	3
	M	99	97	97	97	96	92	86	71	41	48	19	7	4	1	1	5	3	1	2

Note. Percentages are rounded to the nearest whole number. These data are the complement of those obtained for selection of Fill77. Dashes = 0.

quantify the range indexing the size of the low-fill class, the value of the highest fill stimulus that occasioned selection of Fill23 on at least 87.5% of the trials was subtracted from the endpoint value (Fill23). For cases in which no other sample fill value but the endpoint occasioned selection of the endpoint comparison, the range would be 0% and no perceptual class would have been established.

Two-choice primary generalization test performances. As seen in Table 1, no systematic differences were observed in student performances across conditions or comparison selections during two-choice generalization testing. Thus, averaged functions were obtained for students in each group and are depicted in Figure 2. In all cases, a broad range of low-fill stimuli occasioned the nearly exclusive selection of Fill23. With increases in sample fill value, there was a systematic decrease in the likelihood of selecting Fill23. Likewise, a broad range of high-fill stimuli occasioned the nearly exclusive selection of Fill77. With decreases in sample fill value, there was a systematic decrease in the likelihood of selecting Fill77. In addition, because only two choices were used, the gradients used to measure selection of Fill77 and Fill23 were the complements of each other and, thus, did not reveal two independent regions along the fill dimension. Regardless of the number of two-choice trials, the generalization gradients obtained for performances in all conditions overlapped. Thus, the likelihood of selecting Fill23 or Fill77 in the presence of a given fill value did not differ systematically with changes in the number of two-choice generalization test trials.

Forced-choice generalization test data were also analyzed in a 3 × 2 (Number of Two-Choice Trials × Comparison Stimulus Selection) analysis of variance (ANOVA). The range of low- or high-fill stimuli that occasioned the nearly exclusive selection of Fill23 or Fill77 did not differ as a function of the number of test trials, $F(1, 12) = 2.11$, *ns.* In addition, within a group, the range did not differ with selection of Fill23 or Fill77, $F(2, 12) = 4.12$, *ns.* Because forced-choice generalization test performances did not differ systematically, any differences in subsequent three-choice generalization test performances could not be attributed to performance differences during forced-choice testing.

Three-choice generalization test performances. Figure 3 shows the likelihood of selecting Fill77 in the presence of all 19 fill values during three-choice generalization testing as a

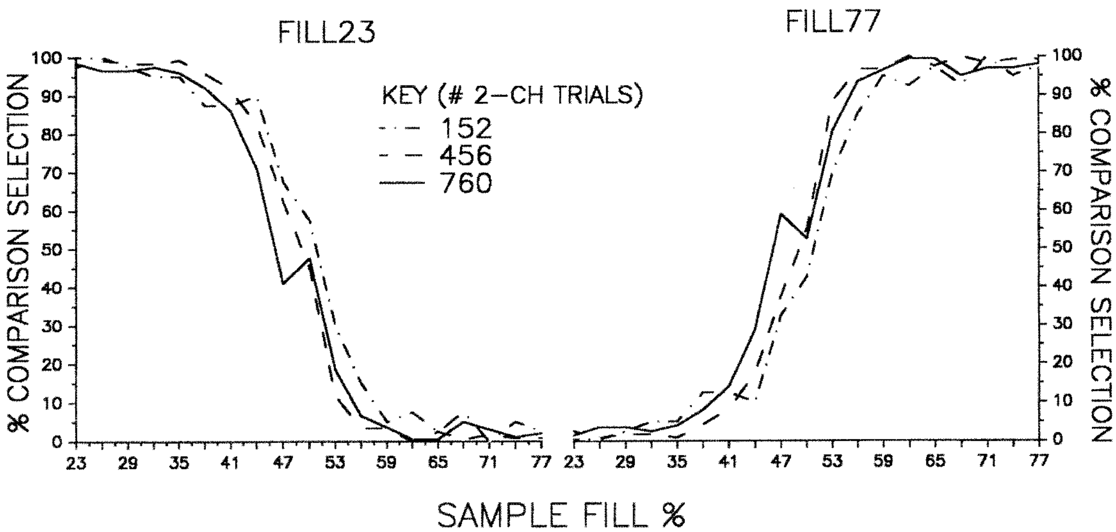


Fig. 2. Two-choice primary generalization test group performances showing the likelihood of selection of Fill23 and Fill77 plotted as a function of number of prior two-choice trials in a group.

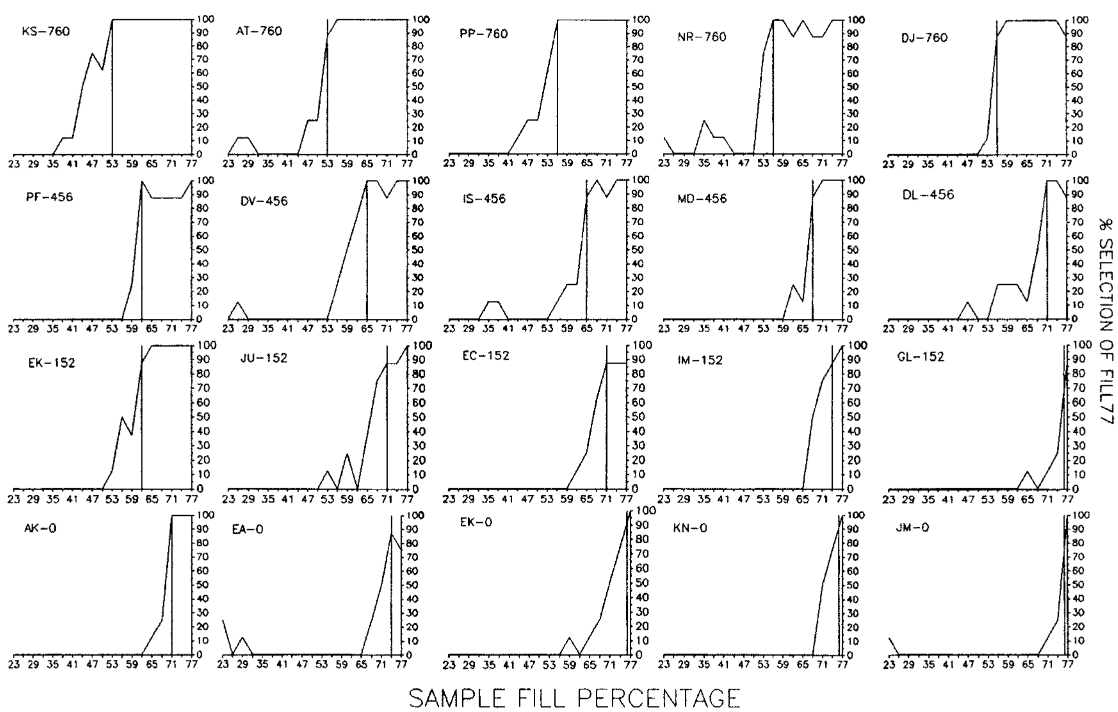


Fig. 3. Individual performances for selection of Fill77 during the three-choice generalization test. Depicted from top to bottom are selection responses from the 760-, 456-, 152-, and 0-trial groups, respectively. Within each condition, performances were ranked from left to right in terms of the width of the range of fill values that occasioned the selection of Fill77 on at least 87.5% of test trials. The range width is delineated by the vertical lines.

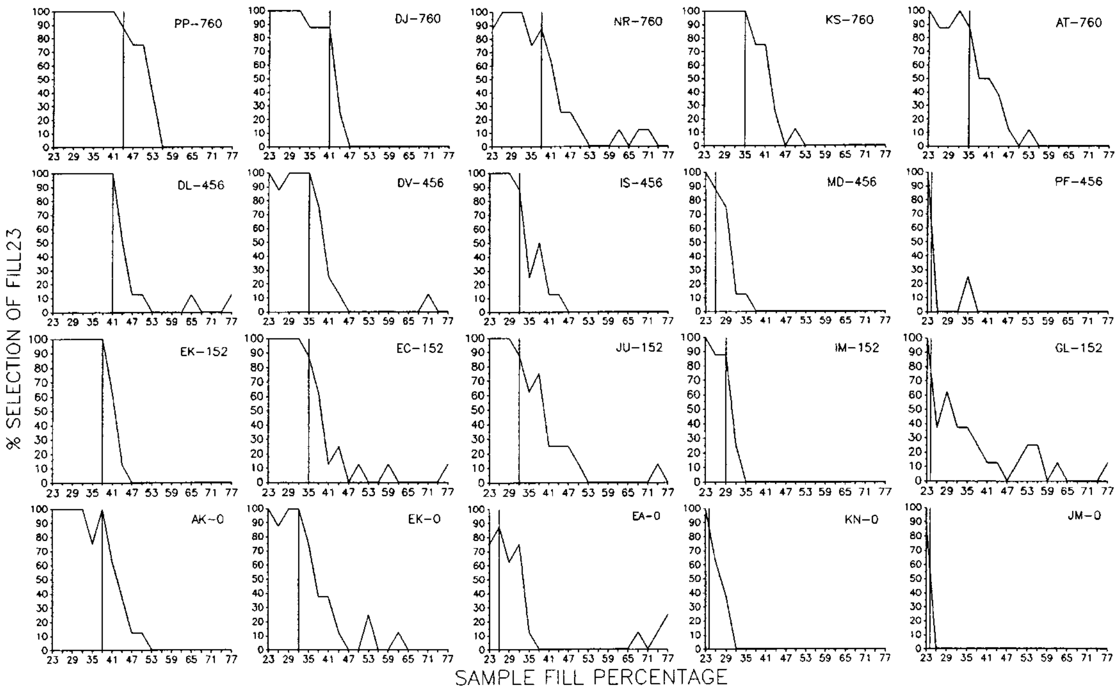


Fig. 4. Individual performances for selection of Fill23 during the three-choice generalization test. The gradients are arranged in the same manner as in Figure 3.

function of the number of prior two-choice generalization trials. From top to bottom, each row represents performances obtained for individual students previously exposed to 760, 456, 152, or 0 two-choice generalization test trials. Students exposed to 760 two-choice trials showed the widest range of high-fill stimuli that occasioned the nearly exclusive selection of Fill77. For the students exposed to 456, 152, and 0 two-choice trials, there was a systematic reduction in this range. Thus, the range of high-fill stimuli that occasioned the nearly exclusive selection of the Fill77 comparison during three-choice generalization testing was a direct function of the number of prior two-choice generalization test trials.

Figure 4 shows the likelihood of selecting Fill23 in the presence of all 19 fill values during three-choice generalization testing. The gradients are arranged in the same manner as in Figure 3. In all cases, the performances of students exposed to 760 two-choice trials showed the widest range of low-fill stimuli that occasioned the nearly exclusive selection of Fill23. Similar reductions in range were produced by prior exposure to both 456 and 152 two-choice test trials. The narrowest range of low-fill stimuli was observed for students who were not exposed to any prior two-choice generalization test trials. Thus, the range of low-fill stimuli that occasioned the nearly exclusive selection of Fill23 during three-choice generalization testing was a direct function of the number of prior two-choice generalization test trials. The effect of prior two-choice testing, however, was not as robust as that observed for the high-fill range (Figure 3).

The three-choice generalization test data depicted in Figures 3 and 4 were also analyzed using a 4 × 2 (Number of Prior Two-Choice Trials × Comparison Stimulus Selection) mixed factorial ANOVA. Collapsing across comparison stimuli selected (Fill23 and Fill77), range was a direct function of the number of prior two-choice generalization test trials, $F(3, 16) = 18.04$, $p < .0001$. Post hoc pairwise comparisons depicted in Table 2 show that the range of high-fill stimuli that occasioned the nearly exclusive selection of Fill77 increased with increases in the number

Table 2

Pairwise comparisons of range differences during three-choice generalization tests in Experiment 1.

	152-high	456-high	760-high	
0-high	−1.55	−3.32**	−7.53**	
152-high		−1.77	−5.98**	
456-high			−4.21**	
	152-low	456-low	760-low	
0-low	−1.55	−1.55	−4.21**	
152-low		0.00	−2.66*	
456-low			−2.66*	
	0-low	152-low	456-low	760-low
0-high	0.88			
152-high		0.89		
456-high			−0.89	
760-high				−2.44*

Note. Low and high indicate the range of stimuli that occasioned selection of Fill23 and Fill77, respectively.
* $p < .05$.
** $p < .01$.

Table 3

Number of students in each experimental group who showed classes of narrow (N), medium (M), and wide (W) ranges for both high- and low-fill classes.

	Class and criterion ranges					
	High			Low		
Group	N (0%–6%)	M (9%–15%)	W (18%–24)	N (0%–6%)	M (9%–15%)	W (18%–24%)
0	5	0	0	4	1	0
152	4	1	0	2	3	0
456	1	4	0	2	2	1
760	0	0	5	0	3	2

Note. These values were derived from the vertical line markers for individual participant performances in Figures 3 and 4.

of prior two-choice test trials. A similar effect was observed when considering the range of low-fill stimuli that occasioned the nearly exclusive selection of Fill23. In addition, for a given number of two-choice trials, there was no systematic difference in the range of low- and high-fill stimuli that occasioned selection of Fill23 and Fill77, respectively, on at least 87.5% of the test trials, $F(1, 16) = 0.60$, *ns.* Finally, range was not influenced by an interaction between the number of prior two-choice test trials and comparison stimulus selection, $F(3, 16) = 2.56$, *ns.*

Although not depicted in Figures 3 and 4, intermediate fill stimuli sometimes occasioned either (a) the exclusive selection of the neither comparison, (b) the complementary selection of the neither comparison and Fill23 but not Fill77, or (c) the complementary selection of the neither comparison and Fill77 but not Fill23. Thus, the inclusion of the neither comparison showed that the two perceptual classes measured by selection of Fill23 and Fill77 comprised functionally independent regions along the fill dimension (Adams et al., 1993; Fields et al., 1993, 1997).

Population effects of forced-choice generalization testing on perceptual class width. As indexed by the range data obtained during three-choice generalization testing, prior forced-choice testing resulted in the formation of potential perceptual classes of different widths. In an additional analysis of this effect, we determined the effects of prior forced-choice testing on the likelihood of the formation of classes of three experimenter-defined widths. A *narrow* class width was defined by the selection of a corresponding endpoint value occasioned by at least the endpoint value and up to two adjacent fill values (range, 0% to 6%). A *medium* class width was defined by the corresponding endpoint plus three to five adjacent fill values (range, 9% to 15%). A *wide* class width was defined by the corresponding endpoint plus at least six adjacent fill values (range, 18% to 24%). This analysis allowed us to determine whether different numbers of prior forced-choice test trials would influence the proportion of students in a group who showed the formation of classes of three different sizes.

Table 3 shows the number of students in a given group who showed potential high-fill classes that were narrow, medium, or wide. As the number of prior forced-choice test trials increased, there were increases in both (a) the modal range of stimuli that functioned as members of the high-fill class and (b) the proportion of students in a group who showed the emergence of these wide classes. The right side of Table 3 shows a similar effect for the low-fill classes, albeit not as robust.

When considering both sides of Table 3, ranges of test fill stimuli defined as medium and wide perceptual classes were observed for 11 of 15 (73%) students exposed to some amount of prior two-choice generalization

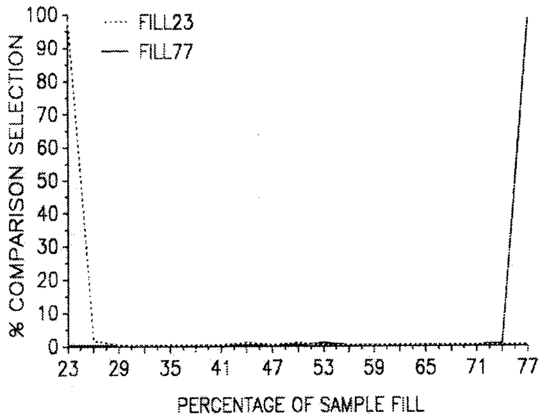


Fig. 5. Mean discriminability test functions for selection of Fill23 and Fill77. The left and right functions represent the likelihood of selecting Fill23 or Fill77, respectively.

testing (Groups 152, 456, and 760). Thus, the majority of the students exposed to two-choice testing showed the emergence of relatively large perceptual classes. In contrast, when considering students who were exposed to identity training but no prior two-choice test trials, only 1 of 5 (20%) showed the emergence of classes that were defined as medium but no larger. This numerical difference (73% vs. 20%) in the likelihood of the emergence of relatively large classes was statistically significant, $\chi^2(1) = 4.44$, $p < .05$.

Discriminability performances. A minimum of 266 trials was required to complete discriminability training and testing. Across students, a mean of 467.4 trials was needed to complete the discriminability training and testing (feedback levels of 100%, 75%, 25%, and 0%). High levels of stimulus control were observed during the first few training trials and were maintained during all subsequent training and testing blocks. Because there were no systematic differences in the postdiscrimination training gradients across the different groups, these data were averaged and are depicted in Figure 5.

Figure 5 shows the likelihood of selecting the Fill23 and Fill77 comparisons for all 19 samples during the last four blocks of discriminability testing. When Fill23 was the sample, the Fill23 comparison was selected nearly exclusively on all trials but was almost never selected in the presence of the 18 other increasingly higher fill values. When Fill77 was the sample, the Fill77 comparison was selected nearly exclusively on all trials but was almost never selected in the presence of increasingly lower fill values. These intermediate fill values (Fill26 through Fill74) occasioned the nearly exclusive selection of the neither comparison. Thus, Fill23 and Fill77 were discriminable from each other and from all other fill values.

Discussion

If specific conditional selection responses occasioned by a small subset of stimuli are also occasioned by a larger range of stimuli without explicit reinforcement, all of these stimuli may be functioning as members of a perceptual class (Adams et al., 1993; Bhatt et al., 1988; Cook et al., 1990; Fields & Reeve, 2000; Fields et al., 1997; Herrnstein, 1990; Keller & Schoenfeld, 1950; Lea, 1984; Wasserman et al., 1988; Zentall et al., 1990). During conditional identity training, students were trained to select the Fill23 comparison when Fill23 was the sample and Fill77 as a comparison when Fill77 was the sample. During the subsequent three-choice generalization test, some range of low-fill stimuli occasioned the nearly exclusive selection of Fill23, and some range of high-fill stimuli occasioned the nearly exclusive selection of Fill77, all without explicitly programmed reinforcement. Thus, these ranges of stimuli may have been functioning as members of perceptual classes.

If these ranges of stimuli are functioning as members of a perceptual class, some of the stimuli within the set must be discriminable from one another (Cook et al., 1990; Fields & Reeve, 2000; Fields et al., 1997; Lea, 1984; Wasserman et al., 1988). Results of the discriminability test showed that Fill23 and Fill77 were discriminable from each other and from all intermediate fill values. Because these performances were established early in discriminability training and were maintained throughout testing, it is likely that discriminability among fill values was present during prior three-choice generalization testing. If so, the range data obtained during three-choice generalization testing did not reflect an inability to discriminate among the fill values. Rather, fill values in a given range evoked the same selection response even though the fill values were discriminable from one another. Thus, these stimuli were functioning as members of two distinct perceptual classes (Bourne et al., 1979; Cook et al., 1990; Fields et al., 1996, 1997; Lea & Ryan, 1984).

Prior to three-choice generalization testing, 5 students received identity training alone (Group 0). In contrast, 15 students received identity training followed by exposure to different numbers of forced-choice generalization test trials (Groups 152, 456, and 760). During the subsequent three-choice generalization test, medium and wide ranges of test fill stimuli occasioned selection of Fill23 or Fill77 for 11 of 15 students (73%) who had been exposed to prior two-choice generalization testing. Because the majority of the students in these groups showed the emergence of perceptual classes, exposure to many forced-choice generalization test trials following identity training reliably predicted the establishment of perceptual classes. In contrast, only 1 of 5 students (20%) who were not exposed to forced-choice generalization test trials showed the emergence of classes that were medium but no larger. Thus, identity training alone was not a reliable predictor of the formation of fill-based perceptual classes. These results suggest that prior exposure to two-choice generalization testing was responsible for the formation of perceptual classes despite the fact that the two-choice test trials did not include concurrently programmed reinforcement.

Although the explicitly programmed reinforcement used to establish identity conditional discriminations was not sufficient by itself to reliably induce the formation of classes, its use most likely had a remote effect on the formation of perceptual classes. Specifically, identity conditional discrimination training established intradimensional discriminative control along the fill dimension (Balsam, 1988). Once intradimensional stimulus control was established, it influenced the conditional selections emitted during the forced-choice generalization tests. Specifically, students were predisposed to select Fill23 in the presence of the lower fill values and to select Fill77 in the presence of the higher fill values. As a result, these similarity-based conditional selections produced a correlated pairing of many low-fill sample stimuli with the Fill23 comparison and many high-fill sample stimuli with the Fill77 comparison, all without reinforcement. These unreinforced pairings, then, most likely established relations among the low-fill stimuli and among the high-fill stimuli. Support for this notion is provided by studies that have demonstrated the establishment of stimulus–stimulus relations without the use of reinforcement (Adams et al., 1999; Leader et al., 1996, 2000; Saunders et al., 1988).

During two-choice generalization testing, the unreinforced conditional selection of an endpoint stimulus in the presence of fill values that were similar to the endpoint stimulus was controlled primarily by the physical similarity of those fill values to the endpoint stimulus. In contrast, the selection of a given endpoint by fill values that were more disparate was controlled to a greater extent by the forced-choice contingencies imposed by the trial format and less by physical similarity. Thus, during forced-choice testing, the combined effect of these two variables resulted in broad generalization gradients. During the subsequent three-choice generalization test, however, the neither response option permitted measurement of the separate effects of both variables.

These separate effects can be illustrated as follows: When 152 trials were presented during the two-choice generalization test, relatively few opportunities were provided for correlated pairings between distal fill values and the endpoint stimuli. This may have precluded the establishment of new stimulus–stimulus relations among these stimuli. Thus, when students in Group 152 were exposed to the three-choice generalization test, the narrow range measures obtained reflected both the relatively weak relations that had been induced between distal fill stimuli and the endpoint stimuli as well as the predominant control of the forced-choice contingencies. As a result, narrow perceptual classes were established.

In contrast, when many forced-choice trials were presented (e.g., 456 and 760), the increased frequency of selections of the endpoint stimulus in the presence of many fill values likely established unreinforced conditional relations between wider ranges of distal fill stimuli and the endpoint. During these two-choice tests, control of the conditional selection of the endpoint stimuli by the distal fill stimuli shifted from the forced-choice contingencies to the new relations established between the distal fill stimuli and the endpoint stimuli. This shift of control increased with increases in the number of prior two-choice trials, as evidenced by the range data obtained during three-choice testing for students in Groups 456 and 760. When the three-choice tests were conducted, then, the selection of the endpoints in the presence of the many more distal fill stimuli reflected the induction of relations among distal fill stimuli and the endpoint as a result of stimulus–stimulus pairings. As a result, increasingly wide perceptual classes were established.

To summarize, the formation of perceptual classes in Experiment 1 can be accounted for by a consideration of three behavioral processes. First, identity conditional discrimination training conducted with reinforcement induced intradimensional control of conditional selections. Second, intradimensional control led to the similarity-based unreinforced conditional selections during forced-choice testing. Finally, with repetition, the selection of endpoint stimuli in the presence of an increasing range of distal stimuli led to the development of relations between the distal stimuli and the endpoint values. This was evidenced by the range measure obtained during three-choice generalization testing, which showed that perceptual class width was influenced by number of prior forced-choice generalization test trials.

EXPERIMENT 2

One purpose of Experiment 2 was to determine whether perceptual classes established by forced-choice testing would also function as a transfer network. To accomplish this, two fill-based perceptual classes were established by exposure to 760 two-choice generalization test trials as in Experiment 1. Students were then trained to select two different glyphs in the presence of Fill23 and Fill77. Finally, a three-choice generalization test determined whether the same range of fill stimuli that occasioned selection of Fill23 or Fill77 also occasioned selection of the glyph related to that fill value through conditional discrimination training.

In Experiment 1, discriminability functions showed that Fill23 and Fill77 were discriminable from all other fill values that comprised each low- and high-fill perceptual class. There was no measure, however, demonstrating whether other members of a perceptual class were discriminable from one another. The second purpose of Experiment 2 was to assess the discriminability of many adjacent stimuli in each fill-based class by way of the presentation of additional discriminability tests.

METHOD

Participants, Apparatus, and Stimuli

Five undergraduate students at Queens College who were unfamiliar with the research area were recruited from introductory psychology classes. The students received partial course credit upon completion of the experiment, but credit was not contingent upon performance. Each student completed the experiment in a single session that lasted approximately 4 to 5 hr.

The apparatus and the 19 fill values used were identical to those in Experiment 1. In addition, during part of Experiment 2 two different black glyphs on a white background (5 cm by 5 cm) served as comparisons: ¶ (Glyph1) and £ (Glyph2).

Procedure

The trial format, block structure, response contingencies, feedback reduction, and responses within a trial were the same as those used in Experiment 1.

Phases 1 through 5. These phases were identical to those for Group 760 in Experiment 1. Included were keyboard familiarization, two-choice identity conditional discrimination training, two-choice primary generalization testing using 760 trials, three-choice fill–fill generalization testing, and training the use of the neither comparison.

Phase 6: Two-choice fill–glyph conditional discrimination training. Next, students received two-choice conditional discrimination training in which either Fill23 or Fill77 served as the sample. The two glyphs served as comparisons. When Fill23 was the sample, the correct response was the selection of Glyph1. When Fill77 was the sample, the correct response was the selection of Glyph2. The neither comparison was not available.

Phase 7: Three-choice fill–glyph generalization testing. Following the establishment of the fill–glyph conditional discriminations, a gen-

Table 4

Percentage of intervals in which Fill23 was selected as a function of sample fill value during the two-choice generalization test in Experiment 2.

Student	Sample fill values																		
	23	26	29	32	35	38	41	44	47	50	53	56	59	62	65	68	71	74	77
ET	98	100	98	100	98	98	95	78	25	28	5	—	—	—	—	—	3	—	—
JC	100	100	100	95	98	95	100	95	70	83	10	8	—	—	—	3	—	—	3
RA	100	100	98	100	100	95	98	83	70	68	20	5	—	—	—	3	—	—	—
MP	100	95	100	100	100	100	93	90	30	40	—	3	3	—	5	—	—	—	3
CF	98	95	98	100	93	98	95	85	67	53	28	5	3	3	—	—	3	3	—
M	99	98	99	99	98	97	96	86	53	54	13	4	1	1	1	1	1	1	1

Note. These performances are the complement of those for selection of Fill77. Percentages are rounded to the nearest whole number. Dashes = 0.

eralization test was presented in which each of the 19 fill values (Fill23 through Fill77) was used as a sample in a random order. Glyph1, Glyph2, and the neither comparison served as comparisons on all trials. Four blocks of 38 trials were presented. Within a block, each glyph comparison was presented an equal number of times on the left and right of the screen in a random order. All comparison selections were followed by noninformative feedback. Phase 7 determined the extent to which generalization gradients obtained during fill–fill generalization testing overlapped with those obtained during fill–glyph generalization testing.

Phases 8 through 12: Discriminability training and testing. Discriminability training and testing were conducted in five phases. Phase 8 was identical to the procedure used in Experiment 1. The 19 fill values served as samples; Fill23, Fill77, and neither served as comparisons. In each of the subsequent four phases of discriminability training and testing, (a) a new pair of fill values served as comparisons along with the neither comparison, and (b) the range of fill samples presented was decreased as follows: In Phase 9, 17 fill values (Fill26 through Fill74) were used as samples along with Fill26, Fill74, and neither as comparisons. In Phase 10, 13 fill values (Fill32 through Fill68) were used as samples along with Fill20, Fill68, and neither as comparisons. In Phase 11, nine fill values (Fill38 through Fill62) were used as samples along with Fill38, Fill62, and neither as comparisons. In Phase 12, five fill values (Fill44 through Fill56) were used as samples along with Fill44, Fill56, and neither as comparisons.

Results

Performances during identity training and two-choice primary generalization testing. All students learned the identity conditional discriminations within a mean of 96.0 trials. In addition, two-choice generalization test performances were similar to those obtained in Experiment 1. The results of two-choice generalization testing are provided in Table 4.

Fill–fill and fill–glyph three-choice generalization test performances. Table 5 provides individual performances during both the fill–fill and fill–glyph three-choice generalization tests conducted in Phases 5 and 7, respectively. These data were used to quantify the range of fill values that occasioned the selection of the Fill23, Fill77, Glyph1 and Glyph2 comparisons on at least 87.5% of the three-choice generalization trials for each individual student. These performances are presented in Figure 6. The top graph shows that a similar range of low-fill samples occasioned selection of Fill23 and Glyph1 on at least 87.5% of the trials for a given student. Although similar within a given student, the range varied across students. The bottom graph shows similar results; a similar range of high-fill samples occasioned selection of Fill77 and Glyph2. In this case, however, the range was identical across all 5 students. Thus, the two ranges of stimuli that occasioned the selection of Fill23 and Fill77 in the presence of all the fill stimuli during three-choice generalization testing were highly predictive of the corresponding ranges of stimuli that occasioned the selection of Glyph1 and Glyph2 during the fill–glyph generalization test.

A 2 × 2 (Fill Comparison Selection ×

Table 5

Individual performances during Phase 5 (fill–fill) and Phase 7 (fill–glyph) three-choice generalization tests in Experiment 2.

Student	Sample fill values																		
	23	26	29	32	35	38	41	44	47	50	53	56	59	62	65	68	71	74	77
Percentage selection of Fill23																			
ET	100	100	100	100	75	75	13	13	—	—	—	—	—	—	—	—	—	—	—
JC	100	100	100	100	100	100	100	88	75	38	13	—	—	—	—	—	—	—	—
RA	100	100	100	100	100	100	63	50	50	50	—	—	—	—	—	—	—	—	—
MP	100	100	100	100	100	88	75	63	50	38	13	—	13	—	—	—	—	—	—
CF	100	100	100	100	100	100	100	100	88	75	50	—	38	—	13	13	13	—	—
M	100	100	100	100	95	93	70	63	53	40	15	0	10	0	3	3	3	0	0
Percentage selection of Glyph1																			
ET	75	100	88	88	75	75	38	38	13	—	—	—	—	—	—	—	—	—	—
JC	100	100	100	100	100	88	100	100	63	50	25	13	13	13	13	—	25	—	—
RA	100	100	88	88	88	63	63	25	—	—	—	—	—	—	—	13	—	—	—
MP	88	100	88	100	88	75	50	38	13	38	13	—	—	—	—	13	—	—	—
CF	88	88	100	100	100	88	75	88	50	38	25	25	—	—	13	—	—	—	—
M	90	98	93	95	90	78	65	58	28	25	13	8	3	3	5	5	5	0	0
Percentage selection of Fill77																			
ET	—	—	—	—	—	—	—	—	25	13	50	63	100	100	100	100	100	100	100
JC	—	13	—	—	—	—	—	—	—	—	13	13	38	88	100	100	100	100	100
RA	—	—	—	—	—	—	—	—	—	—	50	50	50	88	88	100	100	100	100
MP	—	—	—	—	—	—	—	—	13	—	25	38	88	100	100	100	100	100	100
CF	—	—	—	—	—	—	—	—	13	—	—	13	13	88	75	88	88	100	100
M	0	3	0	0	0	0	0	0	10	3	28	35	58	93	93	98	98	100	100
Percentage selection of Glyph2																			
ET	—	—	–	—	13	—	—	13	—	38	38	38	88	100	100	100	100	100	100
JC	—	—	—	—	—	—	—	—	13	—	13	—	13	75	88	88	75	100	100
RA	—	—	—	—	13	13	—	25	—	—	63	63	75	100	100	88	100	100	100
MP	13	13	—	—	—	13	25	13	—	—	—	—	13	75	75	100	100	100	100
CF	13	—	—	—	—	—	—	13	38	—	13	38	75	88	88	88	100	100	100
M	5	3	0	0	5	5	5	13	10	8	25	28	53	88	90	93	95	100	100

Note. Percentages are rounded to the nearest whole number. Dashes = 0.

Glyph Comparison Selection) ANOVA showed that there was no significant difference in the ranges used to measure selection of a fill comparison, $F(1, 8) = 0.83$, *ns*, or a glyph comparison, $F(1, 8) = 0.02$, *ns*. In addition, range was not influenced by an interaction between these two variables, $F(1, 8) = 0.02$, *ns*.

Discriminability performances. Individual performances of all 5 students were similar during the last four blocks of each of the five discriminability tests. Therefore, data were averaged and are depicted in Figure 7, which shows the likelihood of selecting the endpoint fill comparisons in the presence of the fill samples for each of the five discriminability tests. When a low-fill endpoint was the sample, its corresponding low-fill endpoint comparison was selected nearly exclusively on all trials. The low-fill endpoint comparison was rarely selected in the presence of higher sample fill percentages. Although not depicted, these intermediate fill values occasioned the nearly exclusive selection of the neither comparison. The discriminability functions obtained for the selection of the high-fill endpoint comparisons mirror those obtained for selection of the low-fill endpoint comparisons. Based on these data, many of the adjacent stimuli along the dimension of fill were discriminable from one another.

Discussion

During the fill–fill and fill–glyph three-choice generalization tests, a similar range of low-fill stimuli occasioned selection of Fill23 and Glyph1 and a similar range of high-fill stimuli occasioned selection of Fill77 and Glyph2. Subsequent tests of discriminability demonstrated that adjacent fill values within these ranges were discriminable from each other. Thus, prior performances obtained

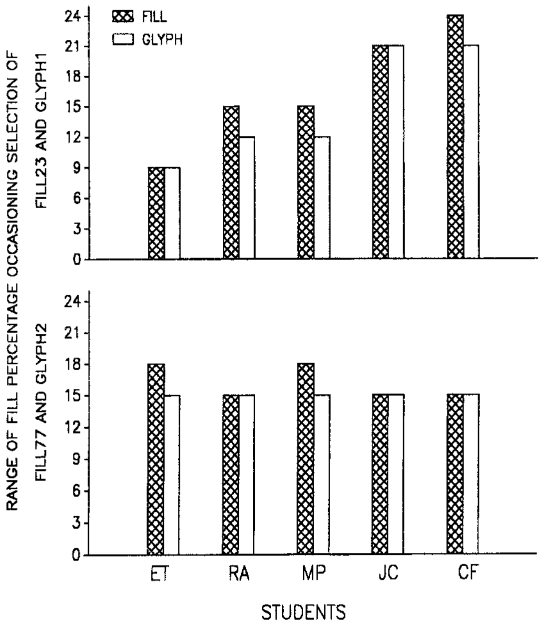


Fig. 6. Range of stimuli that occasioned the nearly exclusive selection of Fill23, Fill77, Glyph1, and Glyph2 for each individual student.

during the fill–fill and fill–glyph three-choice generalization tests did not reflect a failure to discriminate among the fill patterns (Bourne et al., 1979; Cook et al., 1990; Fields et al., 1996, 1997; Lea & Ryan, 1984). Rather, the discriminability of stimuli defined by the range measures demonstrated that the stimuli were functioning as members of perceptual classes.

When a new selection-based response (of a given glyph) was trained in the presence of one member from each of the fill-based perceptual classes, that same response was also occasioned by the remaining class members without benefit of direct training. Thus, the degree of stimulus control exerted by the test fill stimuli was the same when the comparison was physically similar to the sample fill values and when the comparison was a glyph that bore no physical resemblance to the sample fill values. This demonstrates that the perceptual classes established by exposure to forced-choice testing functioned as transfer networks (Dougher et al., 1994; Dougher & Markham, 1994; Fields et al., 1996; Hayes, 1991; Sidman et al., 1989). Taken together with prior studies, the results of Experiment 2 demonstrate that transfer of behavioral function appears to be independent of the mode used for class induction.

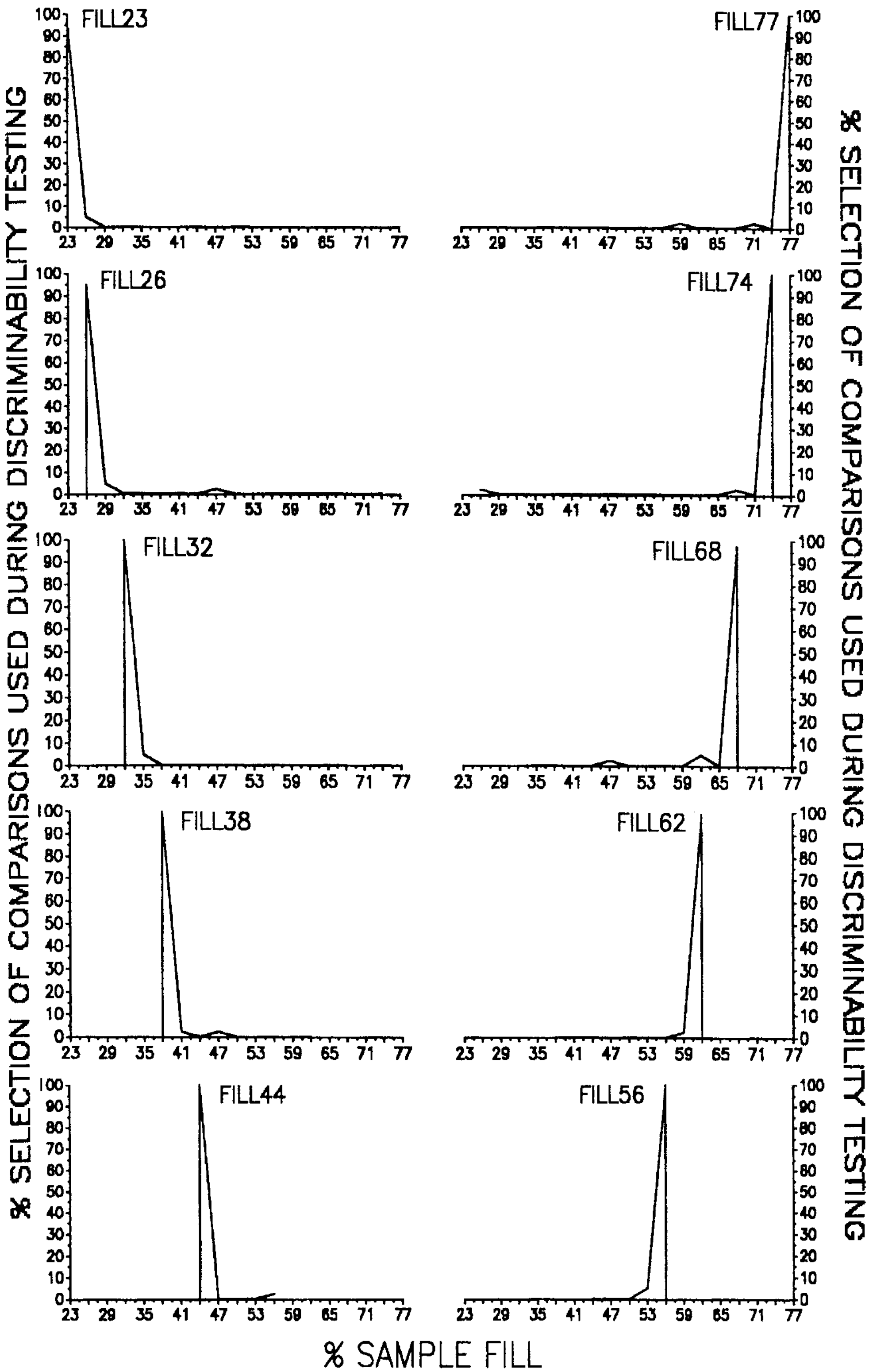
GENERAL DISCUSSION

Traditional methods used to establish perceptual classes have typically involved direct discrimination training using many exemplars drawn from a potential class (e.g., Bhatt et al., 1988; Cook et al., 1990; Herrnstein, 1990; Herrnstein et al., 1976; Honig & Stewart, 1988; Lubow, 1974; Malott & Siddall, 1972; Porter & Neuringer, 1985; Wasserman et al., 1988). The results of Experiments 1 and 2, however, demonstrated that perceptual classes can be established by the presentation of forced-choice generalization testing following intradimensional training with single, rather than multiple, exemplars. It was found that the size of the resulting classes was directly related to the number of prior two-choice generalization test trials. These findings expand the range of operations that can be used to establish perceptual classes and to assess their formation.

Once established through forced-choice testing, the fill-based perceptual classes functioned as transfer networks for the selection of a glyph linked by training to one fill-based perceptual class member. Because the selection of a particular glyph was occasioned by all of the fill values in a given range, these stimuli collectively functioned as a stimulus class. Some stimuli in the class were physically similar (the fill stimuli), and one stimulus (the glyph) was physically disparate from all the other class members. Therefore, the fill–glyph test performances demonstrated that each perceptual class extended beyond its initial defining domain along the fill dimension.

→

Fig. 7. Mean performances for selection of each endpoint value during the last four blocks of each of the five discriminability tests. In the first row, the left and right functions represent the likelihood of selecting Fill23 or Fill77, respectively. In subsequent rows, the left and right functions represent the likelihood of selecting Fill 26 or Fill74, Fill32 or Fill68, Fill38 or Fill62, and Fill44 or Fill56, respectively.


FILL23
FILL77
FILL26
FILL74
FILL32
FILL68
FILL38
FILL62
FILL44
FILL56
% SELECTION OF COMPARISONS USED DURING DISCRIMINABILITY TESTING
% SAMPLE FILL

Because this extended class consisted of some stimuli that were perceptually similar and a stimulus that was perceptually disparate, the class bears formal similarity to other complex stimulus classes, including superordinate semantic categories (Medin & Smith, 1984; Rosch & Mervis, 1975), natural kinds (Gelman, 1988a, 1988b; Gelman & Markman, 1986, 1987) and generalized equivalence classes (Adams et al., 1993; Barnes & Keenan, 1993; Bush, 1993; DeGrandpre, Bickel, & Higgins, 1992; Fields et al., 1991, 1993, 1996; Haring, Breen, & Laitinen, 1989). Thus, all of these names for complex stimulus classes may simply denote the same emergent behavioral phenomenon that was demonstrated in the present studies.

The manner in which the extended fill–glyph classes were established can be illustrated by a consideration of the low-fill extended class. When Fill23 was selected during the forced-choice generalization test, a relation was established between each of the test fill values and Fill23 as a result of unreinforced conditional selections of comparison stimuli (Saunders et al., 1988). These conditional relations can be represented as Fill$(x,y,z) \rightarrow$ Fill23. Next, the Fill23 $\rightarrow$ Glyph1 relation was established through direct conditional discrimination training. Thus, all relations among these stimuli can be represented as Fill$(x,y,z) \rightarrow$ Fill23 $\rightarrow$ Glyph1. Because the different test fill stimuli, the Fill23 endpoint, and Glyph1 were all discriminable from one another, we can label the established relations as A $\rightarrow$ B $\rightarrow$ C.

In the presence of the different test fill stimuli, the likelihood of selecting Glyph1 (C) was high for a range of low-fill stimuli (A). This A $\rightarrow$ C relation, which occurred without direct training, bears similarity to the formal mathematically defined transitive relations that emerge as a result of conditional discrimination training used to establish equivalence classes (Fields & Verhave, 1987). In the case of the Fill$(x,y,z) \rightarrow$ Glyph1 "transitive" relation, Fill23 functions as a nodal stimulus (B) linking each test fill value to Glyph1 (Fields, Adams, Verhave, & Newman, 1990; Fields & Verhave, 1987; Kennedy, Itkonen, & Lindquist, 1994; Spencer & Chase, 1996). Because multiple fill values that are physically similar to one another occasioned selection of Glyph1, taken as a whole these relations bear similarity to a generalized transitive relation. This is one emergent relation necessary to infer the establishment of a generalized equivalence class (Fields & Reeve, 2000). To draw the conclusion that the fill–glyph classes were complete generalized equivalence classes, however, it is necessary to demonstrate the generalization of symmetrical, transitive, and equivalence relations among the stimuli in such a set (Adams et al., 1993; Fields et al., 1996, 1997). Because Experiment 2 included only a demonstration of generalization of transitivity, the inference that generalized equivalence classes were established awaits additional experimentation.

REFERENCES

Adams, B. J., Fields, L., & Verhave, T. (1993). Formation of generalized equivalence classes. *The Psychological Record, 43,* 553–566.

Adams, B. J., Fields, L., & Verhave, T. (1999). Effects of unreinforced conditional selection training, multiple negative comparison training, and feedback on equivalence class formation. *The Psychological Record, 49,* 685–702.

Balsam, P. D. (1988). Selection, representation, and equivalence of controlling stimuli. In R. C. Atkinson, R. J. Herrnstein, G. Lindzey, & R. D. Luce (Eds.), *Stevens' handbook of experimental psychology: Vol. 2. Learning and cognition* (pp. 111–166). New York: Wiley.

Barnes, D., & Keenan, M. (1993). A transfer of functions through derived arbitrary and nonarbitrary stimulus relations. *Journal of the Experimental Analysis of Behavior, 59,* 61–82.

Bhatt, R. S., Wasserman, E. A., Reynolds, W. F., Jr., & Knauss, K. S. (1988). Conceptual behavior in pigeons: Categorization of both familiar and novel examples from four classes of natural and artificial stimuli. *Journal of Experimental Psychology: Animal Behavior Processes, 3,* 219–234.

Bourne, L. E., Dominowski, R. L., & Loftus, E. F. (1979). *Cognitive processes.* Englewood Cliffs, NJ: Prentice Hall.

Bush, K. B. (1993). Stimulus equivalence and cross-modal transfer. *The Psychological Record, 43,* 567–584.

Cook, R. G., Wright, A. A., & Kendrick, D. F. (1990). Visual categorization by pigeons. In M. L. Commons, R. J. Herrnstein, S. Kosslyn, & D. Mumford (Eds.), *Quantitative analyses of behavior: Vol. 8. Behavioral approaches to pattern recognition and concept formation* (pp. 187–214). Hillsdale, NJ: Erlbaum.

Cross, D. V., & Lane, H. (1962). On the discriminative control of concurrent responses: The relations among response frequency, latency, and topography in auditory generalization. *Journal of the Experimental Analysis of Behavior, 5,* 487–496.

DeGrandpre, R. J., Bickel, W. K., & Higgins, S. T. (1992). Emergent equivalence relations between interoceptive (drug) and exteroceptive (visual) stimuli. *Journal of the Experimental Analysis of Behavior, 58,* 9–18.

Dougher, M. J., Augustson, E., Markham, M. R., Greenway, D. E., & Wulfert, E. (1994). The transfer of emotional respondent eliciting and extinction functions through stimulus equivalence classes. *Journal of the Experimental Analysis of Behavior, 62,* 331–351.

Dougher, M. J., & Markham, M. R. (1994). Stimulus equivalence, functional equivalence and the transfer of function. In S. C. Hayes, L. J. Hayes, M. Sato, & K. Ono (Eds.), *Behavior analysis of language and cognition* (pp. 71–90). Reno, NV: Context Press.

Engelmann, S., & Carnine, D. (1982). *Theory of instruction: Principles and applications*. New York: Irvington.

Fields, L., Adams, B. J., Brown, J. L., & Verhave, T. (1993). The generalization of emergent relations in equivalence classes: Stimulus substitutability. *The Psychological Record, 43,* 235–254.

Fields, L., Adams, B. J., Buffington, D. M., Yang, W., & Verhave, T. (1996). Response transfer between stimuli in generalized equivalence classes: A model for the establishment of natural kind and fuzzy superordinate categories. *The Psychological Record, 46,* 665–684.

Fields, L., Adams, B. J., Verhave, T., & Newman, S. (1990). The effects of nodality on the formation of equivalence classes. *Journal of the Experimental Analysis of Behavior, 53,* 345–358.

Fields, L., & Reeve, K. F. (2000). Synthesizing equivalence classes and natural categories from perceptual and relational classes. In J. C. Leslie & D. Blackman (Eds.), *Issues in experimental and applied analyses of human behavior* (pp. 59–84). Reno, NV: Context Press.

Fields, L., Reeve, K. F., Adams, B. J., Brown, J. L., & Verhave, T. (1997). Predicting the extension of equivalence classes from primary generalization gradients: The merger of equivalence classes and perceptual classes. *Journal of the Experimental Analysis of Behavior, 68,* 67–92.

Fields, L., Reeve, K. F., Adams, B. J., & Verhave, T. (1991). The generalization of equivalence relations: A model for natural categories. *Journal of the Experimental Analysis of Behavior, 55,* 305–312.

Fields, L., & Verhave, T. (1987). The structure of equivalence classes. *Journal of the Experimental Analysis of Behavior, 49,* 317–332.

Gelman, S. A. (1988a). Children's expectations concerning natural kind categories. *Human Development, 31,* 28–34.

Gelman, S. A. (1988b). The development of induction within natural kind and artificial categories. *Cognitive Psychology, 20,* 65–95.

Gelman, S. A., & Markman, E. M. (1986). Categories and induction in young children. *Cognition, 23,* 183–209.

Gelman, S. A., & Markman, E. M. (1987). Young children's inductions from natural kinds: The role of categories and appearances. *Child Development, 58,* 1532–1541.

Haring, T. G., Breen, C. G., & Laitinen, R. E. (1989). Stimulus class formation and concept learning: Establishment of within- and between-set generalization and transitive relationships via conditional discrimination procedures. *Journal of the Experimental Analysis of Behavior, 52,* 13–26.

Hayes, S. C. (1991). A relational control theory of stimulus equivalence. In L. J. Hayes & P. N. Chase (Eds.), *Dialogues on verbal behavior* (pp. 19–40). Reno, NV: Context Press.

Hearst, E. (1988). Fundamentals of learning and conditioning. In R. C. Atkinson, R. J. Herrnstein, G. Lindzey, & R. D. Luce (Eds.), *Stevens' handbook of experimental psychology: Vol. 2. Learning and cognition* (pp. 3–110). New York: Wiley.

Herrnstein, R. J. (1990). Levels of stimulus control: A functional approach. *Cognition, 37,* 133–166.

Herrnstein, R. J., Loveland, D. H., & Cable, C. (1976). Natural concepts in pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 4,* 285–301.

Homa, D., & Little, J. (1985). The abstraction and long-term retention of ill-defined categories by children. *Bulletin of the Psychonomic Society, 23,* 325–328.

Honig, W. K., & Stewart, K. E. (1988). Pigeons can discriminate locations presented in pictures. *Journal of the Experimental Analysis of Behavior, 50,* 541–551.

Hrycenko, O., & Harwood, D. W. (1980). Judgments of shape similarity in the Barbary dove (*Streptopelia risoria*). *Animal Behavior, 58,* 586–592.

Innis, A., Lane, S. D., Miller, E. R., & Critchfield, T. S. (1998). Stimulus equivalence: Effects of a default-response option on emergence of untrained stimulus relations. *Journal of the Experimental Analysis of Behavior, 70,* 87–102.

Keller, F. S., & Schoenfeld, W. N. (1950). *Principles of psychology*. New York: Appleton-Century-Crofts.

Kennedy, C. L., Itkonen, T., & Lindquist, K. (1994). Nodality effects during equivalence class formation: An extension to sight-word reading and concept development. *Journal of Applied Behavior Analysis, 27,* 673–683.

Lea, S. E. G. (1984). In what sense do pigeons learn concepts? In H. L. Roitblatt, T. G. Bever, & H. S. Terrace (Eds.), *Animal cognition* (pp. 263–276). Hillsdale, NJ: Erlbaum.

Lea, S. E. G., & Ryan, C. M. E. (1984). Feature analysis of pigeons' acquisition of concept discrimination. In H. L. Roitblatt, T. G. Bever, & H. S. Terrace (Eds.), *Animal cognition* (pp. 233–261). Hillsdale, NJ: Erlbaum.

Leader, G., Barnes, D., & Smeets, P. M. (1996). Establishing equivalence relations using a respondent-type training procedure. *The Psychological Record, 46,* 685–706.

Leader, G., Barnes-Holmes, D., & Smeets, P. M. (2000). Establishing equivalence relations using a respondent-type training procedure III. *The Psychological Record, 50,* 63–78.

Lubow, R. E. (1974). High-order concept formation in the pigeon. *Journal of the Experimental Analysis of Behavior, 21,* 475–483.

Malott, R., & Siddall, J. W. (1972). Acquisition of the people concept in the pigeon. *Psychological Reports, 31,* 3–13.

Medin, D. L., & Smith, E. E. (1984). Concepts and concept formation. *Annual Reviews of Psychology, 35,* 113–138.

Migler, B., & Millenson, J. R. (1969). Analysis of response rates during stimulus generalization. *Journal of the Experimental Analysis of Behavior, 12,* 81–87.

Neisser, U. (1967). *Cognitive psychology.* New York: Appleton Century Crofts.

Njegovan, M., Ito, S., Mewhort, D., & Weisman, R. (1995). Classification of frequencies into ranges by songbirds and humans. *Journal of Experimental Psychology: Animal Behavior Processes, 21,* 33–42.

Omohundro, J. (1981). Recognition vs. classification of ill-defined category exemplars. *Memory & Cognition, 9,* 324–331.

Porter, D., & Neuringer, A. (1985). Music discrimination by pigeons. *Journal of Experimental Psychology: Animal Behavior Processes, 10,* 138–148.

Risley, T. (1964). Generalization gradients following two-response discrimination training. *Journal of the Experimental Analysis of Behavior, 7,* 199–204.

Roche, B., & Barnes, D. (1996). Arbitrarily applicable relational responding and sexual categorization: A critical test of the derived difference relation. *The Psychological Record, 46,* 451–475.

Rosch, E. H., & Mervis, C. B. (1975). Family resemblances: Studies in the internal structure of categories. *Cognitive Psychology, 7,* 573–605.

Rosch, E. H., Mervis, C. B., Gray, W. D., Johnson, D. M., & Boyes-Bream, P. (1976). Basic objects in natural categories. *Cognitive Psychology, 8,* 382–439.

Saunders, R. R., Saunders, K. J., Kirby, K. C., & Spradlin, J. E. (1988). The merger and development of equivalence classes by unreinforced conditional selection of comparison stimuli. *Journal of the Experimental Analysis of Behavior, 50,* 145–162.

Sidman, M., Wynne, C. K., Maguire, R. W., & Barnes, T. (1989). Functional classes and equivalence relations. *Journal of the Experimental Analysis of Behavior, 52,* 261–274.

Spencer, T. J., & Chase, P. N. (1996). Speed analyses of stimulus equivalence. *Journal of the Experimental Analysis of Behavior, 65,* 643–659.

Wasserman, E. A., & DeVolder, C. L. (1993). Similarity and nonsimilarity-based conceptualization in children and pigeons. *The Psychological Record, 43,* 779–794.

Wasserman, E. A., Kiedinger, R. E., & Bhatt, R. S. (1988). Conceptual behavior in pigeons: Categorization of both familiar and novel examples from four classes of natural and artificial stimuli. *Journal of Experimental Psychology: Animal Behavior Processes, 3,* 235–246.

Wittgenstein, L. (1968). *Philosophical investigations* (3rd ed.) (G. E. M. Anscombe, Trans.). New York: Macmillan.

Zentall, T. R., Jackson-Smith, P., & Jagielo, J. A. (1990). Categorical color and shape coding by pigeons. In M. Commons, R. J. Herrnstein, S. M. Kosslyn, & D. B. Mumford (Eds.), *Quantitative analyses of behavior: Vol. 8. Behavioral approaches to pattern recognition and concept formation* (pp. 3–22). Hillsdale, NJ: Erlbaum.

Received October 15, 1998
Final acceptance May 9, 2001